



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Prosodic phenomena

Citation for published version:

Ota, M 2016, Prosodic phenomena: Stress, tone and intonation. in J Lidz, W Snyder & J Pater (eds), *The Oxford Handbook of Developmental Linguistics*. Oxford Handbooks in Linguistics, Oxford University Press, Oxford, pp. 68-86. <https://doi.org/20.500.11820/8a38b205-285d-4859-80b0-046408d5198a>

Digital Object Identifier (DOI):

[20.500.11820/8a38b205-285d-4859-80b0-046408d5198a](https://doi.org/20.500.11820/8a38b205-285d-4859-80b0-046408d5198a)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

The Oxford Handbook of Developmental Linguistics

Publisher Rights Statement:

©Ota, M. (2016). The development of prosody: Stress, tone and intonation. In J. Lidz, W. Snyder, & J. Pater (Eds.), *The Oxford Handbook of Developmental Linguistics*. (Oxford Handbooks in Linguistics). Oxford: Oxford University Press.

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



Prosodic phenomena: Stress, tone and intonation^{*}

Mitsuhiko Ota

*University of Edinburgh***1. Introduction**

This chapter provides an overview of research on the development of prosodic phonology, or the phonological organization beyond segments, particularly at or above the level of the word. The focus will be on three prosodic phenomena that have received much attention in developmental linguistics: namely, stress, tone and intonation. To avoid overlap in coverage, this chapter will not discuss how these prosodic features are related to other areas of learning, such as speech and word segmentation, phonological processes, or morpho-phonological acquisition. The reader is referred to the relevant portion of Chapters XXX, XXX and XXX for these issues.

For each of these prosodic phenomena, the chapter first describes what is known about its course of development, including perceptual precursors in newborns and very young infants, and the subsequent emergence of general and language-specific properties. Next, it presents the outcomes of research that attempts to interpret the developmental observations within the frameworks of metrical stress theory and autosegmental phonology, models that have been central to theoretical research on prosodic phonology in the past few decades. This is followed by a critical assessment of the evidence and arguments for this approach to understanding the development of prosodic phenomena. The chapter concludes with suggestions for future directions.

^{*} To appear in J. Lidz & J. Pater (Eds.), *Oxford Handbook of Developmental Linguistics*. Oxford: OUP. Pre-proof draft. Do not cite without permission.

2. Acquisition of stress

2.1. Development of stress

For the purpose of this chapter, stress is understood as a lexically assigned property of a syllable that renders the syllable a potential position of prominence (Hayes 1995; Sluijter 1995; Ladd 2008). Working from this definition, stress is a structural notion that does not necessarily translate to actual phonetic prominence; the realization of the latter depends on various factors, such as whether the stressed syllable is in or out of the focus position of the utterance. Stress is also a separate matter from the presence and shape of particular pitch patterns (e.g., a high pitch), whose association with a stressed syllable is dictated by the intonational system of the language. Despite the lack of isomorphic phonetic characteristics of stress, an underlyingly stressed syllable can be differentiated from an unstressed one as being phonetically salient in *some* contexts, and it is the acoustic signal of such prominence that the learner must be using to learn the structure and function of stress.

There is much evidence that sensitivity to acoustic differences associated with stress is already present in very young infants. A study using the high-amplitude sucking paradigm shows that newborns can discriminate natural samples of Italian disyllables and trisyllables differing in stress position (e.g., /^lmama/ vs. /ma^lma/, /^ltacala/ vs. /ta^lcala/) (Sansavini, Bertoncini and Giovanelli 1997).¹ Similarly, English-exposed 1-month-olds can detect a change in synthesized disyllables with different stress patterns such as /^lbada/ vs. /ba^lda/, (Spring and Dale 1977, Jusczyk and Thompson 1978). As the stressed syllables in these studies differed from the

¹ A superscript vertical line is placed before the location of primary stress and a subscript vertical line before the location of secondary stress, if any.

unstressed ones in either duration (Sansavini et al. 1997, Spring and Dale 1977) or a combination of duration, amplitude and fundamental frequency (Juszyk and Thompson 1978), the results indicate that neonates are able to differentiate syllables based at least on duration, which is one of the key phonetic correlates of stress in mature systems (Sluijter and van Heuven 1995; Gussenhoven, 2004, Kochanski, Grabe, Coleman, and Rosner 2005).

One of the first signs of language-specific stress development we see in infants is their recognition of the predominant stress pattern of the ambient language. In English, stress frequently falls on the initial syllable of a word (Cutler and Carter 1987), a tendency that is particularly strong in infant-directed speech, in which stress can coincide with the beginning of a word as much as 95% of the time (Kelly and Martin 1994). This characteristic of English stress is picked up by infants some time between 7 and 9 months. Experiments using the head-turn preference procedure show that 9-month-old infants exposed to English (but not those of 6 or 7 months old) prefer to listen to initially-stressed disyllables over finally-stressed disyllables (Juszyk, Cutler, and Redanz 1993; Echols, Crowhurst, and Childers 1997). In German, another language that has an overall tendency toward initial stress, ERP experiments with the mismatch negativity paradigm have shown that a similar bias for initially-stressed disyllables may emerge as early as 4 to 5 months of age (Weber, Hahne, Friedrich, and Friederici 2004, Friederici, Friedrich, and Christophe 2007). The interpretation that such preferences reflect language-specific input rather than a universal bias toward initial stress is reinforced by the lack of comparable effects in infants learning languages without an extremely skewed distribution of stress patterns, such as Spanish and Catalan (Pons and Bosch 2007).

Around the same time, infants' behaviors also begin to reflect crosslinguistic differences in the lexical contrastiveness of stress. While infants respond differently to initially- versus finally-stressed words if they are exposed to a language that uses stress contrastively in lexical items (e.g., English, Spanish, German), they do not if they are learning a language in which stress is not lexically contrastive (e.g., French) (Höhle et al 2009, Pons and Bosch 2010, Skoruppa et al. 2009, 2011). However, there is evidence that infants learning the latter type of language, such as French, do retain sensitivity to acoustic correlates of stress; thus their failure to respond to stress differences imposed on words suggests not a decline in auditory abilities but a functional reorganization due to the non-lexical nature of stress in the language (Skoruppa et al. 2009).

During the first year, infants also gradually become capable of learning stress patterns specific to individual lexical items. In English, the first indication of this process appears in 7-month-olds, who can detect a stress shift in a novel word form they have been familiarized to (e.g., *do'pita* → *'dopita*) (Curtin, Mintz, and Christiansen 2005). Evidence that infants can link such a stress difference to a referential contrast emerges several months later. In experiments using novel word forms and unfamiliar objects, 12-month-olds can learn distinct word-object pairings, even when the word forms differ only in the syllables that are stressed (e.g., *'bedoka* vs. *be'doka*, Curtin 2009, 2011), and 14-month-olds can learn novel pairings when the word forms differ in the position of the stressed syllable (e.g., *'bedoka* vs. *do'beka*, Curtin 2010). Recognition of familiar words by English- or French-exposed 11-month-olds is slowed down when the stress is shifted to the wrong syllable (e.g., *'baby* → *ba'by*), indicating that the representations of real words learned before 1 year

of age already contain information associated with stress (Vihman, Nakai, DePaolis and Hallé 2004).

While the evidence from perception and recognition experiments may suggest that much of stress acquisition is achieved during the first year, production data tell a different story. In spontaneous real-word production as well as nonword imitation, children older than 1 year produce many stress errors, which usually reflect the predominant patterns of the language (English: Kehoe 1997, 1998, 2001, Klein 1984; Dutch: Fikkert 1994, Lohuis-Weber and Zonneveld 1996; Spanish: Hochberg 1988a, 1988b). In English and Dutch, initial primary stress is often imposed on words that begin with a syllable with no stress or secondary stress, as illustrated by the examples in (1):

(1)	<u>Adult target word</u>	<u>Child's production</u>	<u>Age</u>	
	/bə'lɒn/ (Dutch 'balloon')	[ˈmɒmə], [ˈbɒmə]	1;7	(Fikkert 1994)
	/ʃiːˈrɑf/ (Dutch 'giraffe')	[ˈɦɑfə]	2;0	(Fikkert 1994)
	ˌkangaˈroo	[ˈkæŋˌwu]	1;10	(Kehoe 1998)
	/ˌbɛnəˈsiː/ (English nonword)	[ˈbɛˌsiː]	1;10	(Kehoe 1998)

In slightly older children, the conflict between the predominant and more specific stress patterns can result in word productions with two locations of primary stress.

(2)	Adult target word	Child's production	Age
	/bə'lɒn/ (Dutch 'balloon')	[ˈbənˈdɒn]	2;1 (Fikkert 1994)
	/ʃiːˈrɑf/ (Dutch 'giraffe')	[ˈsiːˈaːf]	2;1 (Fikkert 1994)
	ˌkæŋɡəˈruː	[ˈkæŋɡəˈwuə]	2;4 (Kehoe 1998)
	/ˌbɛnəˈsiː/ (English nonword)	[ˈbɛəˈsiː]	2;4 (Kehoe 1998)

Such errors gradually disappear from children's production during the third and fourth years. However, stress patterns involving morphologically complex words and higher levels of prosodic domains continue to undergo development in school-aged children. An example of morphologically conditioned stress patterns would be derivational affixes that induce a predictable stress shift, such as *-ic* and *-ity*, which place primary stress on the preceding syllable (e.g., 'metal → me'tallic, 'personal → perso'nality). These stress shift patterns are not fully acquired by 7- to 9-year-olds (Jarmulowicz 2006). An example of a stress pattern that operates above the level of a simple word is the contrast between compound (a 'hot,dog) and phrasal (a ˌhot 'dog) stress in English. The comprehension of this contrast does not approximate adult performance until children pass the age of 9 years (Atkinson-King 1973, Vogel and Raimy 2002). The later development of these aspects of stress is not surprising since its mastery requires not only purely phonological understanding of stress but also understanding of how stress interacts with morphology (e.g., affixation, compounding) and syntax (e.g., noun phrase structure).

2.2 Metrical phonology and stress acquisition

From the review of the developmental process of stress in the previous section, it should be evident that infants and children do not simply learn stress on an item-by-item basis. They engage in some level of generalization, as indicated by the biases toward language-specific regular patterns in perception and production. This raises questions about the nature of the knowledge of stress in young learners as well as the learning mechanisms involved in the acquisition of stress. The infants' developmental behaviors may emerge from the piecemeal learning of the distributional characteristics observed within numerous instances of individual stress patterns that the learners encounter. Alternatively, the behaviors may be a manifestation of some abstract structural principles that underlie any human-language stress system. A related issue is the extent to which the paths of learning are guided by a priori principles of stress organization. Successful convergence on the adult state may require certain constraints on the range of possible stress systems learners entertain, or it may be sufficiently accomplished through a process that integrates the input data without a predetermined learning space.

In addressing these questions, many researchers have examined the extent to which the development of stress involves the same structural organization of mature phonological systems proposed in metrical stress theory. A central tenet of metrical phonology is that stress reflects a hierarchical structure that governs the positions of prominence in phonological forms (Lieberman and Prince 1977, Hayes 1995, Kager 2007). One common way of representing this underlying structure is through grids of beats that express temporal sequencing along the horizontal axis and prominence along the vertical axis, as in (3).

(3)	(x)	(x)	Word level
	(x .) (x .)	(x .) (x.)	Foot level
	x x x x <x>	x x x x <x>	Syllable level
	L L L L H	L L L H L	
	,hippo'potamus	,memora'bilia	

According to the representation in (3), syllables, potential bearers of stress, are grouped into feet, although for English nouns the final syllable is left out of the grouping (i.e., it is ‘extrametrical’). Feet are composed of a strong position, or its head, (‘x’ at the foot level structure in (3)) and a weak position (‘.’ in (3)). In English, the head of the rightmost foot is the position of word-level prominence (i.e., primary stress). Feet in English are also ‘quantity sensitive’. That is, their formation takes the internal structure of the syllable into consideration. A syllable containing a coda or a long vowel (i.e., a ‘heavy’ syllable, marked H in (3)) can form a foot on its own; otherwise (i.e., if it is ‘light’, marked L in (3)) it needs to be combined with another syllable to form a foot. Metrical representations capture many fundamental characteristics of stress systems in human language, such as the tendency for stress to occur on alternating syllables, and the cumulativeness by which one syllable in a word is singled out to carry the highest prominence.

Metrical representations also allow us to describe crosslinguistic variation systematically. Languages can vary in the type of feet they have in terms of head direction (trochaic or left-headed vs. iambic or right-headed) and quantity sensitivity. Within quantity-sensitive languages, some treat closed syllables as heavy while others do not. Languages also differ in the direction in which syllables are parsed into feet (left-to-right vs. right-to-left), what unit (e.g., syllable, consonant) or position (final or initial) is extrametrical (if any), whether the rightmost or leftmost foot is the most

prominent, and whether feet that do not satisfy their size requirements are still admitted if no other options are available.

On this account, the task of a language learner is to find out the specific pattern of metrical setup adopted by the target language. In a parameter-setting approach (Dresher 1999, Dresher and Kaye 1990), the dimensions of cross-linguistic differences are binary parameters (e.g., parsing direction) that can be set to different values (e.g., right-to-left or left-to-right). The child is modeled as a learner who sets the value of each parameter based on the available cues in the input. In order for the learning to settle on the correct parameter settings, some parameters have been hypothesized to have a default value (the value that is retained in the absence of evidence to the contrary) and the order in which the parameters are set has been prescribed such that the parameters whose values crucially dictates the settings of other parameters are fixed first. In a constraint-based approach, such as Optimality Theory (Prince and Smolensky 1993), stress acquisition has been modeled as algorithmic reranking of violable constraints such as PARSE ('a syllable must be footed') and ALIGN-FEET-RIGHT ('each foot must be aligned with the end of a word'). An example of such a model is Robust Interpretive Parsing/Constraint Demotion (RIP/CD; Tesar 1998, Tesar and Smolensky 2000). In RIP/CD, the stress pattern assigned by the current constraint ranking is compared with the attested pattern. Whenever a mismatch is observed, the algorithm recursively modifies the metrical grammar by demoting certain constraints in the ranking until no such mismatches are detected.² For a fuller discussion of this and other approaches to modeling phonological acquisition with constraints, the reader is referred to Chapter XXX in this volume.

² Tesar (2004) proposes an augmentation (Inconsistency Detection) to this procedure, which allows the learner to resolve ambiguities in the data (i.e., when the attested pattern is consistent with more than one interpretation) by comparing a range of attested forms.

2.3 Evidence for metrical organization in stress development

There are two important empirical questions regarding the proposed involvement of metrical phonology in the development of stress. First, to what extent do patterns in developmental data support the hypothesis that stress learning progresses on the basis of metrical structures? Second, does successful learning of stress necessarily require the structures prescribed by metrical theory and its allied learning models (i.e., parameter-setting, constraint reranking)? This section will address the first question.

In support of the empirical fit between developmental data and metrical phonology, several studies have shown that stages of stress development can be matched up with a systematic progression of metrical settings (Fikkert 1994, Demuth 1996, Kehoe and Stoel-Gammon 1997, Kehoe 1998). For example, Fikkert (1994) bases her account of Dutch stress acquisition on the parameter-setting model of Dresher and Kaye (1990), and presents stage-wise analyses of production data in child Dutch. The stage during which weak-strong targets are produced as strong-weak forms (as illustrated in (1)) is explained as one in which the parameters are set to allow only one left-headed disyllabic foot. The next stage is characterized by level stress (as illustrated in (2)), which exemplifies a metrical stage that allows more than one foot (which is now a left-headed moraic foot), but lacks a setting for the main stress parameter that assigns primary stress.³ Crucially, the unset main stress parameter should result in phonetic forms with level stress, which are not attested in the input. The presence of such forms cannot be directly explained by appealing to convergence to a dominant stress pattern.

³ One problem with this approach to modeling prosodic acquisition is that the putative ‘stages’ of development in reality are not discrete as assumed in the models, but rather overlap in time. This issue has been addressed in a more recent simulation of the same Dutch data using a probabilistic, gradual learning model of constraint reranking (i.e., Boersma’s (1998) Gradual Learning Algorithm) (Curtin and Zuraw 2002).

A related but slightly different type of argument uses the size and shape of early word production as a source of evidence for metrical organization of developmental patterns. One robust observation of children's word production is that there is an initial stage where syllables are omitted from long target words in such a way that the resulting word production is limited to two syllables at most. Furthermore, the child form only admits the dominant prosodic pattern of one or two-syllable words in the language (e.g., a strong(-weak) pattern in English). This pattern has been attested in a number of languages including English (Allen and Hawkins 1978, Echols and Newport 1992, Salidis and Johnson 1997, Schwartz and Goffman 1995), Catalan (Prieto 2006), Dutch (Fikkert 1994, Wijnen, Krikhaar and den Os 1994), Hungarian (Fee 1995), K'iche Maya (Pye 1992) and Japanese (Ota 2003b). This stage of development follows from the general developmental hypothesis in Optimality Theory that all markedness constraints are ranked above faithfulness constraints in the initial state (Smolensky 1996, Davidson, Jusczyk and Smolensky 2004). Such a ranking scheme predicts that the faithfulness constraint that prohibits deletion of input materials (MAX) should be outranked by markedness constraints that require every syllable to belong to a foot (PARSE- σ), feet to be either disyllabic or bimoraic (FTBIN) and every foot to be aligned with a prosodic word (ALIGN-FT) (Pater 1997). As demonstrated by the tableau in (4), the result is that no structure larger than a prosodic word consisting of a single binary foot can be an optimal output, hence the disyllabic maximality effect.⁴ Crucially, this argument rests on the assumption that children's words have metrical constituents consisting of binary feet that group syllables into a hierarchical structure.

⁴ The tableau is adapted from Pater (1997). Round brackets indicate footing, and square brackets boundaries of the prosodic word.

(4)

Input: hippopotamus	ALIGN-FT	PARSE- σ	FTBIN	MAX
a. [(híppo)(pòta)mus]	*!	*		
b. [(híppo)(pòta)]	*!			***
c. [(póta)mus]		*!		****
d. [(pótamus)]			*!	****
e. [(pómus)]				*****

Finally, findings that infants can generalize the underlying stress patterns of nonsense words beyond the surface forms of the familiarization stimuli may be interpreted as evidence that stress acquisition involves abstract principles of metrical structures. For instance, Gerken (2004) and Gerken and Boltt (2008) exposed 9-month-olds to 3- and 5-syllable nonsense words derived from two artificial languages, and then tested their response to tokens which have surface stress forms that are different from those of the familiarization items but still in accordance with the stress assignment grammar of each language. Both artificial languages followed two metrical principles: the Weight-to-Stress principle (i.e., heavy syllables are stressed) and avoidance of stress clash (i.e., no two consecutive stressed syllables are stressed), the latter of which took precedence over the former when their demands conflicted. They also assigned stress according to iterative parsing of syllables into disyllabic feet, albeit in two opposite directions. In Language 1, stress fell on alternating syllables beginning with the initial one but only when it did not interfere with the Weight-to-Stress principle or avoidance of stress clash. In Language 2, the alternating pattern started with the final syllable, again subject to the Weight-to-Stress principle and avoidance of stress clash. The crucial test items were the words in (5), where the syllables in capital letters were stressed.

- (5) a. do TON re MI fa
 b. do RE mi TON fa

The stress pattern in (5a) is consistent with the grammar of Language 1 but not with that of Language 2 (which would have generated ‘do TON re mi FA’). Conversely, (5b) is consistent with Language 2 but not with Language 1 (which would have generated ‘DO re mi TON fa’). The listening times for these test items differed depending on which language the infants were familiarized with. The effects were observed not only when the heavy stressed syllables had the same segmental composition in the training and the test items (e.g., TON; Gerken 2004), but also when they were different, as long as the training set contained more than 2 examples of heavy stressed syllables (e.g., BOM, KEER, SHUL; Gerken and Boltt 2008). These results suggest that 9-month-olds are able to generalize beyond the stress pattern in the individual nonsense words to new words and new stress patterns that reflect a metrical system.

However, it has also been argued that many of the observations cited above as evidence for metrical phonological development can also be simulated in neural networks (Gupta and Touretzky 1994, Shultz and Gerken 2005) or exemplar-based models (Daelemans, Gillis and Durieux 1994; Eddington 2000) without any recourse to formal metrical mechanisms. Learning simulations carried out with these computational models resemble real developmental data in important ways. First, the distribution of errors produced by the simulator during the early stages of learning matches that in children’s word production. Eddington’s (2000) simulation of Spanish stress acquisition yielded the largest number of errors for antepenultimate-stressed target words, much fewer for final-stressed targets, and the fewest for penultimate targets, correctly predicting the order of error frequencies in Hochberg’s (1988a, 1988b) experiments with 3- and 4-year-old Spanish-speaking children. Second, these

simulations often exhibit apparent patterns suggestive of metrical organization. In Daeleman et al. (2004), a simulator trained on Dutch words learned to favor stress assignment to heavy syllables, just as predicted by the Weight-to-Stress principle. In fact, the effect was more pronounced in a model using simple phonemic representations of syllables than in a model in which syllables were annotated for weight. Third, the simulations yield results that correspond with the markedness predictions of metrical theory. For example, Dresher and Kaye (1990) propose that the default (thus, unmarked) setting of foot parsing is iterative. The learner can reset the value of this parameter to the marked setting of non-iterative feet when absence of secondary stress is observed. Consistent with this prediction, the simulators in Gupta and Touretzky (1994) took longer to learn languages with non-iterative feet, even though no explicit bias against non-iterative systems was engineered into the models.

2.4 Metrical theory and the learnability of stress

A second question related to the developmental evidence for the involvement of linguistic principles is whether the acquisition of stress necessarily requires a priori knowledge of metrical phonology in order for the learner to always arrive at the correct target system. Metrical theory limits the types of stress systems that are allowed in human language, and such restrictions on the learner's hypothesis space may be necessary for successful learning to occur. One prediction that follows from this model is that stress patterns that are not licensed by principles of metrical theory should be unlearnable. This prediction has been tested by Gerken and Bollt (2008). Recall that in one experiment in this study, 9-month-olds were shown to generalize weight sensitivity when they were familiarized with words that favored placement of stress on closed (and hence 'heavy' in metrical terms) syllables such as BOM, KEER

and SHUL. In another experiment, they familiarized 9-month-olds to a system in which stress was attracted to open syllables with a /t/ onset (TU, TO, TI). In this case, the infants did not generalize this pattern to novel items. By contrast, 7-month-olds who participated in the same experiment did learn the pattern. Given the standard assumption in metrical phonology (e.g., Hayes 1989) that syllable weight is only sensitive to rhyme structure (i.e., syllable minus the onset), this is a rather surprising result. The interpretation offered by Gerken and Boltt (2008) is that ‘constraints on generalization’ are not inherent in the learners, but develop over time as infants become familiar with the input regularities in the ambient input, which in English includes the tendency for closed syllables to attract stress. There is an alternative interpretation of these outcomes, however. Recent research shows that, although rare, some languages do exhibit onset-based distinctions that can be equated with syllable weight (Gordon 2005). Interestingly, in these languages, low sonority onsets (such as /t/, as in the Gerken-Boltt language) tend to attract more stress than high sonority onsets (such as /n/). The onset-based language in Gerken and Boltt (2008), then, might after all be a possible pattern in natural language, and metrical phonology may have to be revised to incorporate this pattern. The results can, therefore, be reinterpreted as a demonstration of 7-month-olds’ readiness to learn this metrical option after limited exposure because they are still not as committed to the specific ambient pattern (i.e., the rhyme-only syllable weight system of English) as 9-month-olds are.

Another way to investigate what successful acquisition of stress is dependent on is through computer simulations. The amount and type of constraints or biases a simulation needs in order to correctly succeed to learn a possible stress pattern or fail to learn an unattested (hence presumed impossible) pattern can tell us what types of

structure must be hard-wired into the learning model. For example, Hayes and Wilson (2008) demonstrate that a learning model using only structurally-adjacent (or local) information cannot succeed in acquiring nonlocal aspects of stress, such as the assignment of main stress to the rightmost stressed syllable, unless it is augmented by metrical representations of the kind illustrated in (3). The critical insight is that metrical representations make long-distance relationships, such as the positions of stressed syllables, local at some level of analysis. Also using model simulations, Pearl (2011) argues that even when learners adopt parametric metrical phonology, they cannot successfully converge on the target stress system through probabilistic learning unless there is some built-in bias in the learning process. In Pearl's analysis, such a bias must guide the learner toward the input data that unambiguously lead to the correct system – a kind of data that turns out to be a small minority of the input data.

Here again, the caveat raised for the connection between developmental data and metrical theory in the previous section may apply. As pointed out by Gupta and Touretzky (1994), the existence of formal architectures such as metrical principles cannot be proven by demonstrating that they turn an otherwise unlearnable attested system into a learnable one because degrees of learnability can also be changed by models that do not impose formal constraints on the learning space. Furthermore, there is a possibility that the child's initial target system is not the same as the adult stress grammar (Pearl 2011). If so, then any demonstration of learnability or lack thereof based on the description of the adult system may not hold true.

In recent years, the issue of learnability has also been readdressed within a programmatic approach that links language acquisition with typology. Instead of asking what needs to be hard-wired into learners for them to successfully converge on the target system based on finite samples of the language, this approach asks whether

basic characteristics of language (such as the range of attested stress systems) can be explained as a product of children's 'analytic bias' (Moreton 2008). Biases in data interpretation will limit the types of generalizations that children make over a sample of stress patterns, and the outcome of that learning will result in stress systems with common properties that reflect those biases. In an attempt to explore this bidirectional relationship between learning and language universals, Heinz (2009) examined the typology of stress systems in Gordon (2002), and noted that, with a few possible exceptions, all systems are 'neighborhood distinct'. Informally put, this means that if words are construed as a string of steady states (i.e., the beginning, the end and any position between two syllables) and transitions (i.e., types of syllables, such as stressed and unstressed), no two steady states have exactly the same transition pattern before and after them (for a technical definition of neighborhood distinctness, refer to Heinz 2009). By postulating a learner who only learns neighborhood distinct systems, we can explain why children can arrive at least at a *typologically* possible stress system when there are many more *logically* possible systems that account for the examples that they may encounter. In turn, the restriction on the learning can explain why stress systems in human language share some basic characteristics.

3. Tone and intonation

This section begins by discussing tone and intonation together, as the distinction between the two is a matter of linguistic function. In general terms, tone refers to the linguistic use of pitch in marking lexical items, and intonation refers to non-lexical use of pitch to indicate, for example, utterance level pragmatic distinctions (statement vs. question) and phrase boundaries. All languages are known to have intonation, and some languages (about 60-70% of the world's languages, according to Yip (2002)) also have tonal marking of lexical items. Languages that are tonal can differ in how densely they specify the pitch configuration of lexical items, ranging from nearly every syllable (or mora), as in some so-called 'lexical tone languages' such as Mandarin and Yoruba, to only one location per word, as in so-called 'pitch accent languages' such as Japanese, Serbo-Croatian and Swedish.

3.1 Sensitivity to pitch as a linguistic phenomenon

The main phonetic feature of tone and intonation is pitch, a psychophysical correlate of fundamental frequency (F_0). From birth, infants exhibit sensitivity to fundamental frequencies in nonlinguistic stimuli, and can discriminate pure tones that differ only in F_0 (Wormith, Pankhurst and Moffitt 1975). But the perception of pitch cannot be straightforwardly equated with that of F_0 . On one hand, adult listeners perceive sounds that share the same F_0 as having the same tonal height regardless of the composition of the harmonic overtones. On the other hand, if tonal complexes contain harmonics derived from a single F_0 , they are perceived as having the same pitch even when the F_0 itself is missing from the signal. Research using operant conditioning has shown that both of these basic properties of pitch perception are present by 7 months (Clarkson and Clifton 1985, Montgomery and Clarkson 1997). By contrast, 4-month-

olds are incapable of extracting pitch from the combination of overtones, indicating that the perception of pitch is still under development during the first several months (Bundy, Columbo and Singer 1982).

Sensitivity to pitch differences in early infancy has also been demonstrated with linguistic stimuli. Nazzi, Floccia and Bertoncini (1998) used the high-amplitude sucking procedure to show that newborns in France can discriminate two lists of disyllabic Japanese words differing in F_0 contour (ascending vs. descending). Experiments using synthesized speech stimuli show that by 1–2 months, infants can also discriminate contour differences within a syllable (Morse 1972, Kuhl and Miller 1982).

During the same period of development, there also appears to be a more general underlying neural change that differentiates responses to pitch differences that are linguistically relevant from those that are not. Using near-infrared spectroscopy with Japanese-learning infants, Sato, Sogabe and Mazuka (2010) measured hemodynamic brain responses to falling vs. rising pitch contours in pure tone and also in words. At 4 months, responses were bilateral for both the pure tone and word form stimuli, but at 10 months, responses were stronger in the left hemisphere only when the infants heard the contours embedded in words.

3.2 Development of tone

Some perceptual reorganization with respect to linguistic pitch occurs between 6 and 9 months of age. In experiments using the headturn preference paradigm or the stimulus alternating preference procedure, infants learning English, French or Mandarin all respond to the rising versus low tone in Thai at 6 months (Mattock and Burnham 2006, Mattock, Molnar, Polka and Burnham 2008). But at 9 months, only

the Mandarin-learning infants demonstrate sensitivity to this tonal difference. Similarly, 6- to 8-month-old Yoruba-learning infants attend more to F_0 differences among monosyllables than do their English-learning counterparts (Harrison 2000). These findings suggest that infants exposed to lexical tone languages maintain a higher degree of sensitivity to certain types of pitch patterns in comparison to infants learning a language without lexical tone. However, the effect does not appear to be simply caused by a general typological difference in prosodic systems, as English-learning infants continue to show fairly good discrimination of other tonal differences (e.g., Thai rising vs. falling contours). It is more likely that the perceptual difference arises from the phonetic details of the pitch contours that have linguistic functions in the ambient language. For example, the Thai contrast between rising and low tone has some resemblance to the Mandarin tone contrast between rising and low dipping, but such contour difference may not play a major role in English intonation. Conversely, a difference similar to the Thai rising vs. falling contrast signals the difference between the rising and falling intonation in English.

A number of studies have been carried out on the production of tones in languages with lexical tone including Mandarin (Li and Thompson 1977, Clumeck 1980, Hua and Dodd 2000, Wong, Schwartz and Jenkins 2005), Cantonese (Tse 1978, So and Dodd 1995), Taiwanese (Tsay 2001), and Sesotho (Demuth 1993, 1995). Many of these, particularly those studying Asian languages, use transcribed data or adult-listener judgment of spontaneous word production to establish an order of acquisition in the tonal inventory by comparing the (impressionistic) accuracy levels of different tones. Although there is some consistency of order within each language, no apparent universal order of acquisition can be identified among comparable tones (for example, the high level tone is acquired first in Mandarin, but last in Thai). One general

observation that can be made, however, is that tones that are variable in their surface realizations tend to be acquired later than those that are not subject to alternations. Thus, as pointed out by Clumeck (1980), the confusion between the rising tone and low dipping tone in Mandarin-learning children recorded as late as 3½ years can be attributed to the variable realizations of an underlying low dipping tone, which can surface as a rising tone when it follows another dipping tone, or a low level tone if it is elsewhere in a non-final position. In a similar vein, the so-called subject marker tone in Sesotho that invariably differentiates first/second person (no tone) from third person (high tone) is produced fairly accurately by the age of 2 years, while the phonetically similar tonal contrast that underlies verbal roots, highly variable due to its interaction with various phonological processes, is not fully acquired until the age of 3 years (Demuth 1995).

The development of prosodic phonology in languages with a lexical pitch accent system has been studied primarily using spontaneous speech data from Japanese (Hallé, de Boysson-Bardies and Vihman 1991, Ota 2003a) and Swedish (Engstrand, Williams and Strömquist 1991, Kadin and Engstrand 2005, Ota 2006, Peters and Strömquist 1996). These languages feature complex pitch contours that are made up of lexical and intonational components. Studies by Kadin and Engstrand (2005) and Ota (2006) show that the speech production of Swedish children exhibit both components by 18 months, but the complex contours of Japanese children before the age of 18 months sometimes lack a critical intonational feature, i.e., phrase-initial lowering (Ota 2003a). In languages with a lexical pitch accent system, therefore, there is a certain degree of independence between the development of lexical pitch features and intonational pitch features.

3.3 Development of intonation

By 4 to 5 months, infants begin to show evidence for sensitivity to intonational units in speech. In headturn preference experiments, 4.5-month-olds prefer to listen to passages with pauses inserted between clausal boundaries rather than those with the pauses inserted in other places (Jusczyk, Hohne and Mandel 1995). Similar preference for phrasal boundaries appears around 9 months (e.g., Jusczyk, Hirsh-Pasek, Kemler Nelson, Kennedy, Woodward and Piwoz 1992). Further evidence that infants have precocious sensitivity to the global prosodic well-formedness of utterances comes from findings showing that infants as young as 2 months remember a list of words better when they are said with an intonational phrase rather than a prosodically disconnected sequence (Mandel, Jusczyk and Kemler Nelson 1994; Mandel, Kemler Nelson and Jusczyk 1996).

A crucial functional feature of intonation is that variations in pitch patterns signal non-lexical differences. In an extensive review of literature on early spontaneous production, Snow and Balog (2002) conclude that there is no clear evidence that children acquire the form-meaning/context mapping of intonational pitch patterns before the onset of word production. Much of what might be perceived as ‘intonation’ before word production is essentially paralinguistic; in other words, they are modulations of pitch, amplitude and/or speech rate to indicate the emotional states of the speaker, rather than linguistic contrasts. Suggestions have also been made that the pitch contours on early words may be lexically bound (Galligan 1987, Crystal 1979, Halliday, 1975); that is, pitch patterns are learned as if they are part of the lexical property of the word. However, even these studies report productive non-lexical use of pitch from around 17 to 18 months. Children’s understanding of the non-lexical nature of intonational phonology has also been demonstrated in experimental work

using the novel-word/novel-object pairing paradigm. For example, 2½-year-old English-learning children treat novel word forms as different words when they have different vowels, but not when they have different pitch contours (rise-fall vs. low-fall) (Quam and Swingley 2010).

The range of structures and meanings signalled by intonation is quite wide, and not surprisingly, the development of the different functions of intonation is not uniform. Functions that approximate adult-like performance before school age include the use of intonation to differentiate illocutionary acts such as statement versus question (Patel and Grigos 2006), or to mark information structure such as newness (MacWhinney and Bates 1978, Wonnacott and Watson 2008), topic (Chen 2011), and contrastive focus (Hornby and Hass 1970, Wells, Peppé and Goulondris 2004, Müller, Höhle, Schmitz and Weissenborn 2006, Chen 2007). Müller et al. (2006), for example, show that German 4-year-olds produce focus elements ('Peter bakes *a cake*' contrasted with 'Eva wants to bake cookies') with a higher F_0 than non-focused elements. Not all aspects of intonation for information structure are acquired at the same pace, however. For instance, 4- to 5-year-old Dutch-speaking children are capable of producing adult-like contours for topic-marking, but not for focus-marking in the sentence-final position (Chen 2011).

In general, the use of intonation to demarcate the phrasal structure of sentences seems to be acquired later than those functions mentioned above. Comprehension studies show that English-speaking 4- to 5-year-olds fail to reliably use prosodic cues to disambiguate structures such as [Tap][*the frog with the flower*] vs. [Tap *the frog*][*with the flower*] (Snedeker and Trueswell 2004). Although 5- and 7-year-olds understand the syntactic difference between [[*pink and green*] *and white*] vs. [*pink and* [*green and white*]] indicated by intonational phrasing (Beach, Katz and

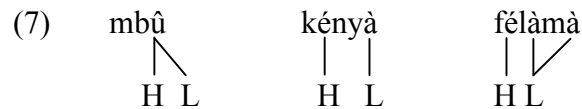
Skowronski 1996), their ability to produce the same prosodic difference has yielded mixed results (Katz, Beach, Jenouri and Verma 1996, Wells et al. 2004).

3.4 Evidence for autosegmental representation in tone and intonation acquisition

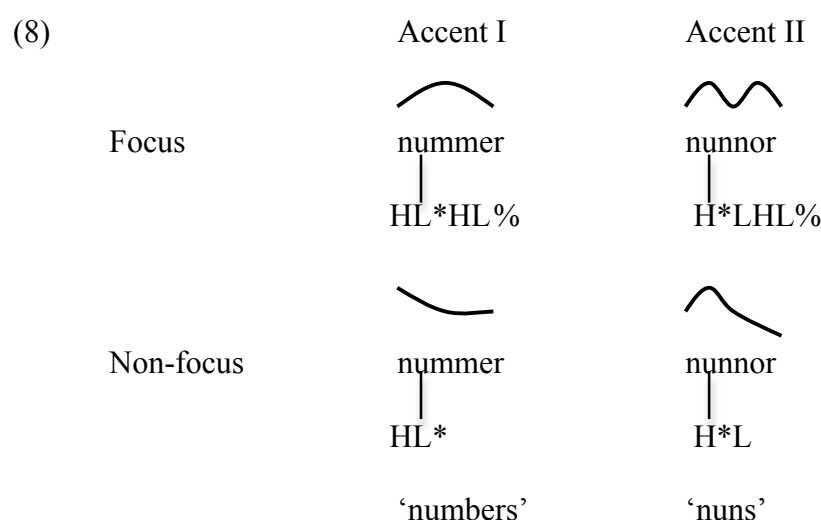
There are two key properties in tone and intonation that may dictate the ways in which they are acquired. First, there is considerable evidence that the phonological elements behind tone and intonation are inherently independent of other phonological structures. Thus, lexical tones can be mobile (e.g., a particular tone can move from one segmental position to another), or stable (e.g., a particular tonal pattern can stay even when the associated segment or syllable is deleted), and participate in one-to-many or many-to-one association with other structures (e.g., a particular tone can be ‘spread’ over many segments or syllables) (Yip 2002). Second, because all tonal and intonational phonology has a single phonetic correlate (i.e., pitch), the mapping between the acoustic signal and the underlying phonology can be complex. A particular pitch configuration not only can be a marker of lexical distinction, phrase boundary or an utterance type, but also may be a composite of all of them.

These properties of tone and intonation have been successfully modeled in autosegmental phonology, which postulates discrete tonal elements lined up along a separate tier from the rest of phonological representation, both for lexical tone (Leben 1975, Liberman 1975, Goldsmith 1976) and intonation (Pierrehumbert 1980, Beckman and Pierrehumbert 1986, Ladd 2008). The independence of the tonal tier from the segmental tier can be illustrated by the following examples from Mende. The three words here differ in the number of syllables as well as the surface contour

of pitch.⁵ Nevertheless, autosegmental representations allow us to see that they have the same underlying tonal structure: H(igh)-L(ow).



As an example of autosegmental analysis of the interaction between different types of tonal and intonational phonology, (8) shows the two ‘word accents’ in Stockholm Swedish, which exhibit variable contour realizations depending on whether or not they appear in a focus position (including their citation forms). This complex pattern can be explained as a combination of lexical pitch accents (either H*L or HL*, where the asterisk indicates the tone that is associated with the stressed syllable), an intonational phrasal accent (H), and an utterance final L% tone lined up on the same tonal tier. The lexical pitch accents are always present, but the phrasal accent occurs only when the word is in a focus position (Bruce 1977, 1987).



⁵ A circumflex indicates a falling tone, an acute accent a level high tone, and a grave accent a level low tone.

One can see how an understanding of pitch-related phonology in terms of atomic tonal units may assist the learning of tone and intonation. If quantitatively continuous F_0 information is abstracted into strings of discrete units, it can constrain the possible phonological structures that can be postulated for the attested data.⁶ Complex contours that may reflect patterns associated with a range of diverse functions including lexical contrasts, phrasal boundaries, or discourse semantics can be decomposed into units of mapping between tonal sequences and their functions. General patterns behind alternations in pitch patterns can be learned as autosegmental processes (e.g., spreading of tonal association, avoidance of identical adjacent tones). In this way, autosegmental phonology provides a plausible model of the acquisition of tone and intonation.

Whether the development of these phenomena actually involves autosegmental mechanisms is, of course, an empirical question. The literature offers several types of supporting evidence. The first type of evidence for autosegmental structure comes from various observations pointing to the separation of pitch patterns from segmentals. If tone is acquired as an inherent feature of vowels, sensitivity for nonnative tonal contrasts should also begin to attenuate around the same time. Infants' sensitivity to nonnative vowel (quality) contrasts typically declines before 6 months (Kuhl, Williams, Lacerda, Stevens and Lindblom 1992, Polka and Werker 1994). However, results in Mattock and Burnham (2006) and Mattock et al. (2008) indicate that the analogous perceptual reorganization of tone takes place later, some time between 6 and 9 months. Such findings suggest that the perceptual development of tones is independent of that of vowel quality.

⁶ How learners can translate the continuous and time-varying acoustic signal of pitch into discrete phonological is not a trivial question. But recent computational work promises a solution (Yu, xxx).

Some non-adult-like pitch patterns found in early production also indicate the separation of tonal and intonational phonology from segmental structures (Demuth 1993, 1995; Ota 2003a). One such example in Demuth's analysis of early Sesotho involves the application of the Obligatory Contour Principle (the ban on adjacent identical tones on the tonal tier). In Sesotho, when two underlying high tones become adjacent, one of them becomes a low tone. In autosegmental terms, delinking of a high tone occurs to satisfy the OCP, and the toneless tone-bearing unit receives a default low tone. The example in (9a) is an utterance produced by a 2½-year-old Sesotho speaker, who omitted the subject marker from the target structure (which is given in 9b) (Demuth 1993: 297).⁷ The omission would have made the high tone in /ná/ adjacent to the high tone in the first syllable of the verb stem (/bídíkìsà/). Instead, the stem-initial syllable was produced with a low tone (/bídíkìsà/), resulting in a structure that respects the OCP. Crucially, the tonal specification of a syllable in the adult model has changed in the child production, indicating the independence of tonal and segmental representations in the child's phonological system.

- (9) a. ná bídíkìsà
 b. nná kè-à-bídíkìsà
 1sPN 1sSM-PRES-turn
 'Me, I'm revolving (it)'

Occasionally, children come up with an idiosyncratic system of phonology that they appear to have spontaneously created. Such original language games or word

⁷ In Demuth's transcription, only high tones are marked (by an acute accent), and all other vowels are assumed to have low tone. For clarity sake, low-tone vowels have been marked here with a grave accent. 1sPN = First-person singular pronoun; 1sSM = First-person singular subject marker; PRES = Present tense.

plays offer unique evidence for autosegmental representations of tonal structures. Yue-Hashimoto (1980), for instance, discusses the case of a Mandarin-speaking child who productively engaged in a word game from the age of two years. The word game involved the application of a fixed tonal pattern to real words, as shown in (10)⁸. All disyllabic words received a HL pattern, regardless of the original tones (10a). Monosyllabic words were reduplicated and also forced into the HL template (10b), unless the nucleus contained two vowels, in which case the vowels were split into separate syllables with an LH pattern (10c). Although data like this may not generalize to ‘normal’ language development, they serve as a striking demonstration of children’s ability to manipulate tonal elements separately from segments.

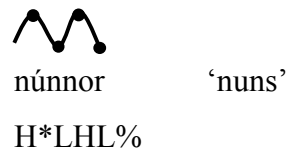
(10)	<u>Adult target word</u>	<u>Word play</u>
a.	t’au ₃₅ wan ₃₅	t’au ₅₅ wan ₁₁
b.	yən ₃₅	yən ₅₅ yən ₁₁
c.	k’ua ₅₃	k’u ₁₁ l a ₅₅

Another type of argument for autosegmental representations in early phonology relates to the idea that pitch movements in (adult) tonal and intonational phonology are best represented as sequences of discrete tonal elements rather than the shape and slopes of the contours. A corollary of this model is that pitch contours tend to have phonetically stable turning points that anchor the pitch movement. For example, the so-called ‘Accent II’ in citation forms of Swedish words (given in (8)) has been analyzed as having 4 underlying tones: H*LHL%. Bruce (1977, 1987) shows that the phonetic constants in such pitch configurations are the F₀ of points that correspond to

⁸ Tones in these examples are marked with the standard number notation used for Asian languages: 35 = ‘high-rise’, 55 = ‘high level’, 11 = ‘low level’.

the L and H of the hypothesized underlying tonal structure (i.e., the black dots superimposed on the contour in (11)).

(11)



Using second-degree polynomials defined by high and low turning points, Ota (2006) examined the spontaneous speech of Swedish-speaking 16- to 18-month-olds, and found that more high-low-high turning point sequences can be identified in Accent II words than in other productions. Furthermore, the higher the F_0 peak of the stressed syllable was, the larger the drop after the peak, demonstrating the relative stability of the low F_0 point, a presumed phonetic realization of the L tone between the two H tones.

4. Summary and future directions

The purpose of this chapter was to review some key descriptive findings in the development of stress, tone and intonation, and to discuss the extent to which the acquisition of these phenomena can be understood in light of the structural representations and formal organizational principles proposed within metrical theory and autosegmental theory.

There now exists a substantial body of descriptive work in this area that provides us with some understanding of how stress, tone and intonation develop over time, from the pre-linguistic sensitivity shown by very young infants to late-acquired

aspects of these prosodic properties. Further progress in this area is contingent on having better developmental data, both in terms of coverage and quality. There is a noticeable lack of work outside the usual stock of familiar languages (e.g., English, Dutch, German, French, Chinese, Japanese). Particularly problematic is the paucity of information on the acquisition of languages with an iambic system of stress or an African type tone system. Without access to developmental data from such systems, it is difficult to test the full range of typological predictions that follow from metrical approaches to prosodic acquisition.

The quality of developmental data also needs to improve. Much of the production data we currently have are based on transcriptions, which are not only difficult to verify, but also prone to adult-listener's bias. This is a particularly serious concern for stress, which has a complex relationship with its phonetic correlates. As we do not fully understand how stress is acoustically signaled in early production, adult transcribers who listen for stress using their phonetic correlates may miscode the data. The issue of phonetic realization also applies to laboratory-based developmental research. For instance, most experimental studies do not fully control for the acoustic dimensions that vary between the stressed and unstressed syllables in the stimuli. As a result, we cannot be sure what cues are used by the participant infants or children to determine the presence or position of stress.⁹

Turning now to more theoretical issues, it still remains an open question whether the acquisition of prosody is best understood within metrical and autosegmental phonology, despite the many attempts to make the case. With respect to the development of stress, we have seen that the affinity between the predictions of metrical theory and developmental data does not in itself warrant a causal relationship

⁹ A recent study that addresses this failing is Lintfert and Möbius (2010).

between them, since at least some of the developmental observations are also consistent with models that do not assume the abstract structures featured in metrical phonology. On the other hand, it is clear that these frameworks provide us with extremely useful representational devices to describe and characterize the developing system and its difference from the adult system. It is an added advantage of the metrical approach that the hypothesized developing stress systems can be subsumed within a general architecture of phonology that has been proposed to account for mature systems. The issues of learnability and its connection to metrical structure spawn empirical questions that can direct our exploration of the potentially inherent mechanisms (constraints or biases) involved in the development of stress. Furthermore, they afford an impetus to examine how language acquisition is related to the typological properties of phonological systems attested in human languages.

Similarly, several types of arguments can be put forward to support the idea that tone and intonation are acquired as sequences of tonal elements linearly organized independently from other phonological structures, as proposed in autosegmental theory. Unlike in research on stress development, however, these claims have not been systematically pitted against developmental models that do not rely on pre-wired structural units. This is probably due to the fact that our understanding of tone and intonation per se (and hence their development) has generally lagged behind that of stress. But this state of affairs is rapidly changing with recent developments in prosodic modeling. Most of the work in this area explores ways to best model pitch contours, using, in some cases, the same discrete and static tonal categories in autosegmental phonology (e.g., ToBI; Silverman et al 1992), but in other cases, more articulatorily or acoustically motivated targets or parameters (e.g., PENTA, Xu and Wang 2001; INTSINT, Hirst and Di Cristo 1998; Tilt, Taylor 2000). Computationally

informed work is also emerging in the area of tonal and intonational acquisition, addressing non-trivial questions such as how tonal categories can be learned from continuous speech signals, and what acoustic information, types of learning functions, and potential structure in the hypothesis space may be necessary for tonal and intonational learning to succeed (Gauthier, Shi and Xu 2007, 2009, Yu 2011). Both types of research are likely to shed new light on the representational requirements and learning mechanisms involved in the development of tone and intonation.

References

- Allen, G. D. and S. Hawkins (1978). 'The development of phonological rhythm'. In A. Bell and J. Hooper (eds.), 173-185.
- Atkinson-King, K. (1973). 'Children's acquisition of phonological stress contrasts'. *UCLA Working Papers in Phonetics* 25.
- Archibald, J. (ed.) (1993). *Phonological acquisition and phonological theory*. Hillsdale, Lawrence Erlbaum.
- Beach, C., W. F. Katz and A. Skowronski (1996). 'Children's processing of prosodic cues for phrasal interpretation'. *Journal of the Acoustical Society of America* 99, 1148-1160.
- Beckman, M. and J. Pierrehumbert (1986). 'Intonational structure in English and Japanese'. *Phonology Yearbook* 3, 255-310.
- Bell, A. and J. Hooper (eds.) (1978). *Syllables and segments*. New York, Elsevier.
- Boersma, P. (1998). *Functional phonology: Formalizing the interactions between articulatory and perceptual drives*. Doctoral dissertation, University of Amsterdam. The Hague: Holland Academic Graphics.

- Bruce, G. (1977). *Swedish word accents in sentence perspective*. Lund, Geerup.
- (1987). ‘How floating is focal accent?’ In K. Gregerson and H. Basbøl (eds.), 41-49.
- Bundy, R. S., J. Columbo and J. Singer (1982). ‘Pitch perception in young infants’. *Developmental Psychology* 18, 10-14.
- Chen (2007). ‘Intonational realization of topic and focus by Dutch-acquiring 4-to-5-year-olds. In *Proceedings of the 16th International Congress of Phonetic Sciences (ICPhS)*, 1553-1556.
- (2011). ‘Tuning information packaging: Intonational realization of topic and focus in child Dutch’. *Journal of Child Language* 38, 1055-1083.
- Clarkson, M. G. and R. K. Clifton (1985). ‘Infant pitch perception: Evidence for responding to pitch categories and the missing fundamental’. *Journal of the Acoustical Society of America* 77, 1521-1527.
- Clumeck, H. (1980). ‘The acquisition of tone’. In G. H. Yeni-Komshian, J. F. Kavanagh and C. A. Ferguson (eds.), 257-275.
- Crystal, D. (1979). ‘Prosodic development’. In P. Fletcher and M. Garman (eds.), 174-197.
- Cutler, A and D. M. Carter (1987) ‘The predominance of strong initial syllables in the English vocabulary’. *Computer Speech & Language*, 2, 133-142.
- Curtin, S. (2009). ‘Twelve-month-olds learn novel word-object pairings differing only in stress pattern’. *Journal of Child Language* 36, 1157-1165.
- (2010). ‘Young infants encode lexical stress in newly encountered words.’ *Journal of Experimental Child Psychology* 105, 376-385.

- T. H. Mintz and M. H. Christiansen (2005). 'Stress changes the representational landscape: Evidence from word segmentation'. *Cognition* 96, 233-262.
- K. Zuraw (2002). 'Explaining constraint demotion in a developing system. *Boston University Conference on Language Development* 26, 118-129.
- Daelemans, W. S., S. Gillis and G. Durieux (1994). 'The acquisition of stress: A data-oriented approach'. *Computational Linguistics* 20, 421-451.
- Davidson, L., P. W. Jusczyk and P. Smolensky (2004). 'The initial and final states: Theoretical implications and experimental explorations of Richness of the base'. In R. Kager, J. Pater and W. Zonneveld (eds.), 321-368.
- de Lacy, P. (2007). *The Cambridge handbook of phonology*. Cambridge: Cambridge University Press.
- Demuth, K. (1993). 'Issues in the acquisition of the Sesotho tonal system'. *Journal of Child Language* 20, 275-301.
- (1995). 'The acquisition of tonal systems'. In J. Archibald (ed.), 111-134.
- (1996). 'The prosodic structure of early words'. In J. L. Morgan and K. Demuth (eds.), 171-184.
- Dresher, E. (1999). 'Charting the learning path: Cues to parameter setting'. *Linguistic Inquiry* 30, 27-67.
- and J. Kaye (1990). 'A computational learning model for metrical phonology'. *Cognition* 34, 137-195.
- Echols, C. and E. Newport (1992). 'The role of stress and position in determining first words'. *Language Acquisition* 2, 189-220.

- Echols, C., M. J. Crowhurst, and J. B. Childers (1997). 'The perception of rhythmic units in speech by infants and adults'. *Journal of Memory and Language* 36, 202-225.
- Eddington, D. (2000). 'Spanish stress assignment within the analogical modeling of language'. *Language* 76, 92-109.
- Engstrand, O, K. Williams and S. Strömqvist (1991). 'Acquisition of Swedish tonal word accent contrast'. In *Proceedings of the 12th International Congress of Phonetic Sciences (ICPhS)*, 324-327.
- Fee, E. J. (1995). 'Two strategies in the acquisition of syllable and word structure'. *Child Language Research Forum* 27, 29-38.
- Fikkert, P. (1994). *On the acquisition of prosodic structure*. Dordrecht, Holland Institute of Generative Linguistics.
- Fletcher, P. and M. Garman (eds.) (1979). *Language acquisition: Studies in first language development*. Cambridge, Cambridge University Press.
- Friederici, A. D., M. Friederich and A. Christophe (2007). 'Brain responses in 4-month-old infants are already language specific'. *Current Biology* 17, 1208-1211.
- Gauthier, B., R. Shi, and Y. Xu (2007). 'Learning phonetic categories by tracking movements'. *Cognition* 103, 185-205.
- (2009). 'Learning prosodic focus from continuous speech input: A neural network exploration'. *Language Learning and Development* 5, 94-114.
- Galligan, R. (1987). 'Intonation with single words: Purposive and grammatical use'. *Journal of Child Language* 14, 1-21.

- Gerken, L. A. (2004). 'Nine-month-olds extract structural principles required for natural language'. *Cognition* 93, B89-B96.
- and A. Bollt (2008). 'Three exemplars allow at least some linguistic generalizations: Implications for generalization mechanisms and constraints'. *Language Learning and Development* 4, 228-248.
- Goldsmith, J. (1976). *Autosegmental phonology*. PhD thesis, MIT. [Published 1979 by New York, Garland Press.]
- (ed.) (1995). *The Handbook of Phonological Theory*. Oxford, Blackwell.
- Gordon, M. (2002). 'A factorial typology of quantity-insensitive stress'. *Natural Language & Linguistic Theory* 20, 491-552.
- (2005). 'A perceptually-driven account of onset-sensitive stress'. *Natural Language & Linguistic Theory* 23, 595-653.
- Gregerson, K. and H. Basbøl (eds.) (1987). *Nordic prosody IV*. Odense, Odense University Press.
- Gupta, P. and D. S. Touretzky (1994). 'Connectionist models and linguistic theory: Investigations of stress systems in language'. *Cognitive Science* 18, 1-50.
- Gussenhoven, C. (2004). *The phonology of tone and intonation*. Cambridge, Cambridge University Press.
- Hallé, P., B. de Boysson-Bardies and M. M. Vihman (1991). 'Beginnings of prosodic organization: Intonation and duration patterns of disyllables produced by French and Japanese infants'. *Language and Speech* 34, 299-318.
- Halliday, M. A. K. (1975). *Learning how to mean: Explorations in the development of language*. London, Edward Arnold.

- Harrison, P. (2000). 'Acquiring the phonology of lexical tone in infancy'. *Lingua* 110, 581-616.
- Hayes, B. (1989). 'Compensatory lengthening in moraic phonology'. *Linguistic Inquiry* 20, 253-306.
- (1995). *Metrical Stress Theory: Principles and Case Studies*. Chicago, University of Chicago Press.
- Heinz, J. (2009). 'On the role of locality in learning stress patterns'. *Phonology* 26, 303-351.
- Hirst, D. J., A. Di Cristo (eds.) (1998). *Intonation Systems: A Survey of Twenty Languages*. Cambridge, Cambridge University Press.
- Hochberg, J. G. (1988a). 'First steps in the acquisition of Spanish stress'. *Journal of Child Language* 15, 273-292.
- (1988b). 'Learning Spanish stress: Developmental and theoretical perspectives'. *Language* 64, 683-706.
- Hornby, P. A. and W. A. Hass (1970). 'Use of contrastive stress by preschool children'. *Journal of Speech and Hearing Research* 13, 395-399.
- Höhle, B., R. Bijeljac-Babic, B. Herold, J. Weissenborn, and T. Nazzi (2009). 'The development of language specific prosodic preferences during the first year of life: Evidence from German and French'. *Infant Behavior and Development* 32, 262-274.
- Hua, Z. and B. Dodd (2000). 'The phonological acquisition of Putonghua (modern standard Chinese)'. *Journal of Child Language* 27, 3-42.

- Jarmulowicz, L. (2006). 'School-aged children's phonological production of derived English words'. *Journal of Speech, Language, and Hearing Research* 49, 294-308.
- Jusczyk, P. W. and Thompson (1978). 'Perception of a phonetic contrast in multisyllabic utterances by two-month-old infants'. *Perception and Psychophysics* 23, 105-109.
- A. Cutler and N. J. Redanz (1993). 'Infants' preference for the predominant stress patterns of English words'. *Child Development* 64, 675-687.
- K. Hirsh-Pasek, D. G. Kemler Nelson, L. Kennedy, A. Woodward and J. Piwoz (1992). 'Perception of acoustic correlates of major phrasal units by young infants'. *Cognitive Psychology* 24, 252-293.
- E. Hohne and D. Mandel (1995). 'Picking up regularities in the sound structure of the native language'. In W. Strange (ed.), 91-119.
- Kadin, G. and O. Engstrand (2005). 'Tonal word accents produced by Swedish 18- and 24-month-olds'. *Proceedings of FONETIK 2005*, 67-70.
- Kager, R., J. Pater and W. Zonneveld (eds.) (2004). *Constraints in phonological acquisition*. Cambridge, Cambridge University Press.
- (2007). 'Feet and metrical stress' In P. de Lacy (ed.), 195-227.
- Katz, W. F., C. M. Beach, K. Jenouri and S. Verma (1996). 'Duration and fundamental frequency correlates of phrase boundaries in productions by children and adults'. *Journal of the Acoustical Society of America* 99, 3179-3191.
- Kehoe, M. (1997). 'Stress error patterns in English-speaking children's word productions'. *Clinical Linguistics and Phonetics* 11, 389-409.

- (1998). 'Support for metrical stress theory in stress acquisition'. *Clinical Linguistics and Phonetics* 12, 1-23.
- (2001). 'Prosodic patterns in children's multisyllabic word productions'. *Language, Speech, and Hearing Services in Schools* 32, 284-294.
- Stoel-Gammon (1997). 'The acquisition of prosodic structure: An investigation of current accounts of children's prosodic development'. *Language* 73, 113-144.
- Kelly, M. H. and S. Martin (1994). 'Domain-general abilities applied to domain-specific tasks: Sensitivity to probabilities in perception, cognition, and language'. *Lingua* 92, 105-140.
- Klein, H. (1984). 'Learning to stress: A case study'. *Journal of Child Language* 11, 375-390.
- Kochanski, G., E. Grabe, J. Coleman and B. Rosner (2005). 'Loudness predicts prominence: Fundamental frequency lends little'. *Journal of the Acoustical Society of America*, 118, 1038-1054.
- Kuhl, P. and J. D. Miller (1982). 'Discrimination of auditory target dimensions in the presence or absence of variation in a second dimension by infants'. *Perception and Psychophysics* 31, 279-292.
- K. A. Williams, F. Lacerda, K. N. Stevens and B. Lindblom (1992). 'Linguistic experience alters phonetic perception in infants by 6 months of age'. *Science* 255, 606-608.
- Ladd, D. R. (2008). *Intonational phonology*. [second edition]. Cambridge, Cambridge University Press.

- Leben, W. (1975). *Suprasegmental phonology*. PhD thesis, MIT. [Published 1980 by New York, Garland Press.]
- Li, C. and S. Thompson (1977). 'The acquisition of tone in Mandarin-speaking children'. *Journal of Child Language* 4, 185-199.
- Lieberman, M. (1975). *The intonational system of English*. Unpublished PhD thesis, MIT.
- A. Prince (1977). 'On stress and linguistic rhythm'. *Linguistic Inquiry* 8, 249-336.
- Lohuis-Weber, H. and W. Zonneveld (1996). 'Phonological acquisition and Dutch word prosody'. *Language Acquisition* 5, 245-283.
- MacWhinney, B. and E. Bates (1978). Sentential devices for conveying givenness and newness: A cross-cultural developmental study. *Journal of Verbal Learning and Verbal Behavior* 17, 539-558.
- Mandel, D. R., P. W. Jusczyk and D. G. Kemler Nelson (1994). 'Does sentential prosody help infants to organize and remember speech information'? *Cognition* 53, 155-180.
- D. G. Kemler Nelson and P. W. Jusczyk (1996). Infants remember the order of words in a spoken sentence. *Cognitive Development* 11, 181-196.
- Mattock, K. and D. Burnham (2006). 'Chinese and English infants' tone perception: Evidence for perceptual reorganization'. *Infancy* 10, 241-265.
- M. Molnar, L. Polka and D. Burnham (2008). 'The developmental course of lexical tone perception in the first year of life'. *Cognition* 106, 1367-1381.

- Montgomery, C. R. and M. G. Clarkson (1997). 'Infants' pitch perception: Masking by low- and high-frequency noises'. *Journal of the Acoustical Society of America* 102, 3665-3672.
- Morgan, J. and K. Demuth (eds.). (1996). *Signal to syntax: Bootstrapping from speech to grammar in early acquisition*. Mahwah, Lawrence Erlbaum.
- Morse, P. A. (1972). 'The discrimination of speech and nonspeech stimuli in early infancy'. *Journal of Experimental Child Psychology* 14, 477-492.
- Müller, A., B. Höhle, M. Schmitz and J. Weissenborn (2006). 'Focus-to-stress alignment in 4 to 5-year-old German-learning children'. In *Proceedings of GALA*, 393-407.
- Nakayama, M. (ed.) (2001). *Issues in East Asian Language Acquisition*. Tokyo, Kuroshio.
- Nazzi, T., C. Floccia and J. Bertoncini (1998). 'Discrimination of pitch contours by neonates'. *Infant Behavior and Development* 21, 779-784.
- Ota, M. (2003a). The development of lexical pitch accent systems : An autosegmental analysis. *Canadian Journal of Linguistics*, 48, 357-383.
- (2003b). *The development of prosodic structure in early words: Continuity, diverge and change*. Amsterdam, John Benjamins.
- (2006). 'Children's production of word accents in Swedish revisited'. *Phonetica* 63, 230-246.
- Patel, R. and M. I. Grigos (2006). 'Acoustic characterization of the question-statement contrast in 4, 7 and 11 year-old children'. *Speech Communication* 48, 1308-1318.

- Pater, J. (1997). 'Minimal violation and phonological development'. *Language Acquisition* 6, 201-253.
- Pearl, L. (2011). 'When unbiased probabilistic learning is not enough: Acquiring a parametric system of metrical phonology'. *Language Acquisition* 18, 87-120.
- Peters, A. and S. Strömquist (1996). 'The role of prosody in the acquisition of morphemes'. In J. Morgan and K. Demuth (eds.), 215-232.
- Pierrehumbert, J. (1980). *The phonetics and phonology of English intonation*. Unpublished PhD thesis, MIT.
- M. Beckman. (1986). *Japanese tone structure*. Cambridge, MIT Press.
- Polka, L. and J. Werker (1994). 'Developmental changes in perception of nonnative vowel contrasts'. *Journal of Experimental Psychology: Human Perception and Performance* 20, 421-435.
- Pons, F., and Bosch, L. (2007). 'The perception of stress patterns by Spanish and Catalan infants. In P. Prieto, J. Mascaró and M. J. Solé (eds.), 199-218.
- (2010). 'Stress pattern preference in Spanish-learning infants: The role of syllable weight'. *Infancy* 15, 223-245.
- Prieto, P. (2006). 'The relevance of metrical information in early prosodic word acquisition: A comparison of Catalan and Spanish'. *Language and Speech* 49, 23-258.
- J. Mascaró and M. J. Solé (eds.) (2007). *Segmental and prosodic issues in (romance) linguistics. Current issues in linguistic theory*. Amsterdam: John Benjamins.

- Prince, A. and P. Smolensky (1993). *Optimality theory: Constraint interaction in generative grammar*. New Brunswick, NJ and Boulder, CO: Rutgers University and University of Colorado.
- Pye, C. (1992). 'The acquisition of K'iche' Maya'. In D. I. Slobin (ed.), 221-308.
- Quam, C. and D. Swingley (2010). 'Phonological knowledge guides two-year-olds' and adults' interpretation of salient pitch contours in word learning'. *Journal of Memory and Language* 62, 135-150.
- Salidis, J. and J. S. Johnson (1997). 'The production of minimal words: A longitudinal case study of phonological development'. *Language Acquisition* 6, 1-36.
- Sansavini, A., J. Bertoncini and G. Giovanelli (1997). 'Newborns discriminate the rhythm of multisyllabic stressed words'. *Developmental Psychology* 33, 3-11.
- Sato, Y., Y. Sogabe and R. Mazuka (2010). 'Development of hemispheric specialization for lexical pitch-accent in Japanese infants'. *Journal of Cognitive Neuroscience* 22, 2503-2513.
- Schwartz, R. G. and L. Goffman (1995). 'Metrical patterns of words and production accuracy'. *Journal of Speech and Hearing Research* 38, 876-888.
- Shultz, T. R. and L. A. Gerken (2005). 'A model of infant learning of word stress.' In *Proceedings of the 27th Annual Conference of the Cognitive Science Society*, 2015-2020.
- Silverman, K., M. Beckman, J. Pitrelli, M. Ostendorf, C. Wightman, P. Price, J. Pierrehumbert and J. Hirschberg (1992). 'ToBI: a standard for labeling English prosody'. *Proceedings of the 1992 International Conference on Spoken Language Processing* 867-870.

- Skoruppa, K., F. Pons, A. Christophe, L. Bosch, E. Dupoux, N. Sebastián-Gallés, R. Alves Limissuri and S. Peperkamp (2009). 'Language-specific stress perception by 9-month-old French and Spanish infants'. *Developmental Science* 12, 914-919.
- Skoruppa, K., A. Cristià, S. Peperkamp and A. Seidl (2011). 'English-learning infants' perception of word stress patterns'. *Journal of the Acoustical Society of America* 130, EL50-EL55.
- Slobin, D. I. (1992). *The crosslinguistic study of language acquisition, Vol. 3*. Hillsdale, Lawrence Erlbaum.
- Sluijter, A. M. C. and V. J. van Heuven (1995). 'Spectral balance as an acoustic correlate of linguistic stress'. *Journal of Acoustical Society of America* 100, 2471-2484.
- Smolensky, P. (1996). 'On the comprehension/production dilemma in child language'. *Linguistic Inquiry* 27, 720-731.
- Snedeker, J. and J. Trueswell (2004). 'The developing constraints on parsing and decisions: The role of lexical biases and referential scenes in child and adult sentence processing. *Cognitive Psychology* 49, 238-299.
- Snow, D. and H. L. Balog (2002). 'Do children produce the melody before words? A review of developmental intonation research. *Lingua*, 112, 1025-1058.
- So, L. K. H. and B. J. Dodd (1995). 'The acquisition of phonology by Cantonese-speaking children'. *Journal of Child Language* 22, 473-495.
- Spring, D. R. and P. Dale (1977). 'Discrimination of linguistic stress in early infancy'. *Journal of Speech and Hearing Research* 20, 224-232.

- Strange, W. (1995). *Speech perception and linguistic experience: Theoretical and methodological issues in cross-language speech research*. Timonium, York Press.
- Taylor, P. A. (2000). 'Analysis and synthesis of intonation using the Tilt model', *Journal of the Acoustical Society of America* 107, 1697-1714.
- Tesar, B. (1998). 'An iterative strategy for language learning'. *Lingua* 104, 131-145.
- (2004). 'Using inconsistency detection to overcome structural ambiguity'. *Linguistic Inquiry* 35 219-253.
- P. Smolensky (2000). *Learnability in Optimality Theory*. Cambridge, MIT Press.
- Tse, J. K. P. (1978). 'Tone acquisition in Cantonese: A longitudinal case study'. *Journal of Child Language* 5, 191-204.
- Tsay, J. (2001). 'Phonetic parameters of tone acquisition in Taiwanese'. In M. Nakayama (ed.), 205-226.
- Vihman, M. M., S. Nakai, R. A. DePaolis, and P. Hallé (2004). 'The role of accentual pattern in early lexical representation'. *Journal of Memory and Language* 50, 336-353.
- Vogel, I. and E. Reimy (2002). 'The acquisition of compound vs. phrasal stress: The role of prosodic constituents'. *Journal of Child Language* 29, 225-250.
- Wannacott, E. and D. G. Watson (2008). Acoustic emphasis in 4-year-olds. *Cognition* 107, 1093-1101.

- Weber, C., A. Hahne, M. Friedrich and A. D. Friederici (2004). 'Discrimination of word stress in early infant perception: Electrophysiological evidence'. *Cognitive Brain Research* 18, 149-161.
- Wells, B., S. Peppé and N. Goulondris (2004). 'Intonation development from five to thirteen'. *Journal of Child Language* 31, 749-778.
- Wijnen, F., E. Krikhaar and E. den Os (1994). 'The (non) realization of unstressed elements in children's utterances: A rhythmic constraint'? *Journal of Child Language* 21, 59-83.
- Wong, P., R. G. Schwartz and J. J. Jenkins (2005). 'Perception and production of lexical tones by 3-year-old, Mandarin-speaking children'. *Journal of Speech, Language, and Hearing Research* 48, 1065-1079.
- Wormith, S. J., D. Pankhurst and A. R. Moffitt (1975). 'Frequency discrimination by young infants'. *Child Development* 46, 272-275.
- Xu, Y. and Q. E. Wang (2001). 'Pitch targets and their realization: Evidence from Mandarin Chinese', *Speech Communication* 33: 319-337.
- Yeni-Komshian, G. H., J. F. Kavanagh and C. A. Ferguson (1980) (eds.). *Child phonology, Vol 1. Production*. New York, Academic Press.
- Yip, M. (2002). *Tone*. Cambridge, Cambridge University Press.
- Yu, K. (2011). *The learnability of tones from the speech signal*. PhD dissertation, University of California, Los Angeles.
- Yue-Hashimoto, A. O.-K. (1980). 'Word play in language acquisition: A Mandarin case'. *Journal of Chinese Linguistics* 8, 181-204.

Zonneveld, W. and D. Nouveau (2004). 'Child word stress competence: An experimental approach'. In R. Kager, J. Pater and W. Zonneveld (eds.), 369-408.